ENGINEERING DEPARTMENT

General Radio Company

CAMBRIDGE, MASSACHUSETTS

Reprint No. A-25 1945

Reprinted from ELECTRONICS, January 1945 Issue January 1945

Audible Audio Distortion

BY

H. H. SCOTT

PRINTED IN U.S.A.

Audible AUDIO DISTORTION

By H. H. SCOTT
General Radio Company
Cambridge, Mass.

DISAGREEMENT between the average non-technical listener and the engineer regarding the excellence of reproduced sound has become almost traditional. The listener's judgment is based upon whether or not he finds the reproduction faithful, or at least pleasing. The engineer, on the other hand, tends to have preconceived ideas based upon various technical characteristics, which may or may not be the factors governing the listener's preference.

Laboratory instruments and techniques represent a means and not an end. It is necessary occasionally to reconsider the results obtained through laboratory measurements, to decide whether or not they are indicative of the actual important performance characteristics of the equipment under test. This is particularly true in cases involving human judgment and psychological or physiological factors.

Quest for Perfect Reproduction

The characteristics of systems for the electrical reproduction of sound can be measured in physical terms to a high degree of precision, and such reproducing systems can be designed to perform with any given degree of excellence. A close approach to perfection will be found in certain types of transmitting and recording equipment, which is necessarily expensive. However, in the design of most audio-frequency equipment, and. in particular, radio receivers, phonographs, and sound picture projectors of the types manufactured in large quantities for home use, economic considerations must frequently take precedence over artistic ideals. In such a design perfection is not expected, and the problem is to provide the best possible results, as judged by the listener, within predetermined price limits.

Perfect reproduction of sound is the will-o'-the-wisp that has been chased through half a century by the phonograph, radio, and motion picture industries. Thirty years ago Edison attempted to demonstrate that there was no noticeable difference between the reproduction of his phonograph and the voice of the living artist. Radio advertisements through the years have acclaimed the tinny squeaks or muffled rumblings which were in vogue at the moment as absolute perfection. Actually, early phonographs and radios were so far from perfection that each successive change seemed like a tremendous improvement. Engineers and scientists, however, have never overlooked the shortcomings of what currently passed for perfect tone, and have anticipated the day when the reproduced sound would be indistinguishable from the original. Today this is a technical possibility, limited only by practical and economic factors.

Practical Considerations

In planning for the post-war period the matter of distortion in sound-reproducing systems should be reviewed carefully, to take best advantage of the existing state of the art and provide the public with the best possible tone quality per dollar expended. A hint as to some of the past difficulties and disagreements lies in the word quality. Our laboratory measurements so far are quantitative only. What the listener hears is qualitative, and the relationship between the two is extremely complex and little understood. About all we can rely upon is the fact that, if the commonly recognized types of distortion are reduced below certain measurable levels, the ear will be satisfied with the result. This procedure can be followed in the design of high-priced studio and transmitting equipment, but in the case of sound-reproducing systems for the home it is not yet commercially practical and may never be, because of price competition.

Distortion is of Three Types

It is unfortunate that no single measurement will define the excellence of sound reproduction. Audio distortion is generally classified into three types—namely, frequency discrimination, harmonic (also called amplitude or non-linear) distortion, and phase distortion. In the past the most advantageous balance among these three characteristics has not been maintained. This is the reason for the almost traditional disagreement among engineers, sales departments, and customers.

Frequency discrimination is easily measured with even a relatively simple oscillator. It met early acceptance in engineering circles as a criterion of quality. Actually it is only one of several important characteristics. The terms harmonic distortion and amplitude distortion are misleading and do not convey an impression of the real seriousness of this type of distortion. Phase distortion is important mainly in long transmission lines and other circuits where time delay occurs. The amount present in the usual home amplifier and loudspeaker system is considerably less important than the other types of distortion, but cannot, of course, be neglected entirely if these others are reduced.

Wide Frequency Response is Not Enough

These are the three types of distortion which the engineer considers and attempts to correlate. How do his measurements correspond with the judgment of the average listener, unprejudiced by

---

From a paper presented at the National Electronics Conference, Chicago, 1944.

Wide frequency response alone is not enough for perfect reproduction of sound. Generation of intermodulation frequencies must be suppressed in order to secure favorable listener reaction. A double-beat oscillator for intermodulation measurements is described

technical knowledge? The poor public acceptance of many so-called high-fidelity systems proves that, even when it is really attained, wide frequency response alone is not the answer. In fact, wide frequency response may be a disadvantage if noise or other forms of distortion are present. Considerable research is needed on the correlation between the various forms of distortion. The only information available is incomplete and often contradictory. In the meantime the engineer can attempt to base his measurements and conclusions on factors at least logically related to the average listener's reactions.

The alterations in music caused by variations from a flat frequency response or by moderate phase shifts in the reproducing equipment are not fundamentally different in character from variations which may exist under actual listening conditions with no electrical reproduction interposed. For instance, an orchestra will sound different when playing in different halls, and still differently again when playing outdoors. The acoustic conditions under which the listener hears the music will cause wide variations from what a microphone placed near the orchestra might pick up. When one hears music outdoors, or through a doorway, or from the back seat in a top balcony, it does not sound distorted in the popular sense of the term, and yet the effective transmission of different frequencies between the orchestra and one's ear may vary tremendously in both amplitude and phase. The ear will accept a large amount of this variation without considering the music as unnatural, even though many of the high or low frequencies may be missing entirely. This is probably one reason why the public has been able to tolerate radio receivers with poor frequency-response characteristics.

Intermodulation Products

What the average listener defines as tone is mainly governed by the frequency-response characteristic. A radio has a high tone, a mellow tone, or a deep tone, depending upon the frequency range and the balance between high and low frequencies. The non-technical person does not consider these variations in tone as distortion.

Electrical or mechanical reproducing systems, however, subject the music to another form of distortion which is unnatural because it is never encountered under conditions where the music is heard without reproduction or reinforcement. This is the poorly named amplitude or harmonic distortion. It is not the actual deviation from the original amplitude relationships which in itself is objectionable. Neither is it in most cases the increase in harmonics which were present in the original music at appreciable amplitudes. Associated with this form of distortion is the generation of many intermodulation products of an amplitude equal to or higher than the generated harmonics and bearing no harmonic or musical relationship to the components of the original sound. The importance of this form of distortion has been generally overlooked because of difficulties of exact measurement and interpretation. Actually this form of distortion is probably the most annoying of all types and warrants considerable further investigation.

It has long been noted that correlation between harmonic measurements and actual listening tests is inconsistent. The production of

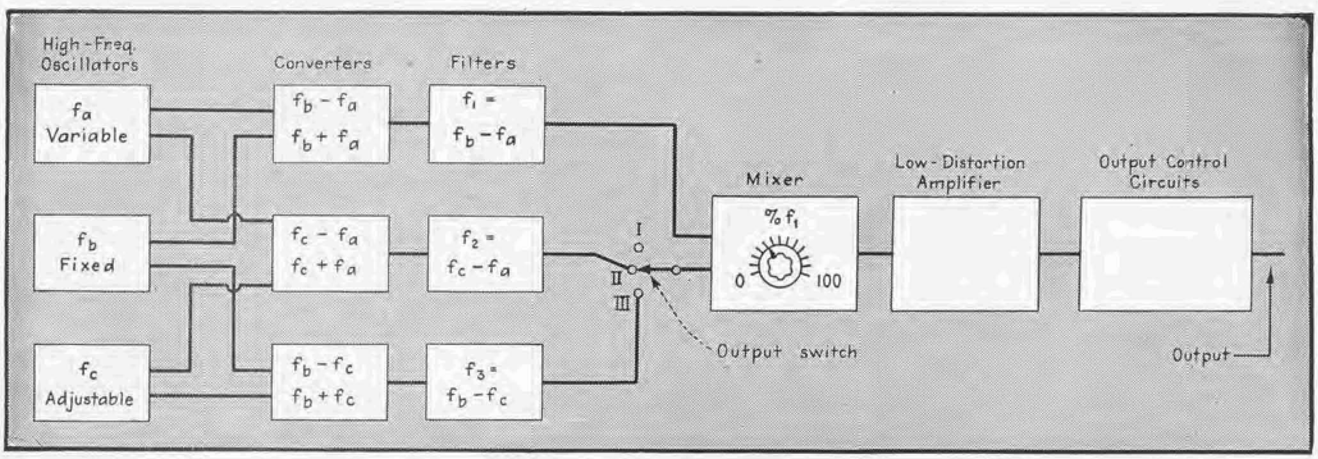

FIG. 1—Principle of operation of double-beat oscillator. At output switch position I the output frequency is f_1 and is variable. At II, both f_1 and f_2 appear at the output, and are so related that $f_1 - f_2 = K$ or $f_1 + f_2 = K$. At III, f_1 and f_3 are the output frequencies, with f_1 variable and f_3 equal to K

intermodulation products is not necessarily proportional to the production of harmonics excepting under certain carefully defined conditions. Hence the conventional methods of measuring harmonic distortion with a harmonic analyzer or distortion meter, which measures the amount of harmonics added to a single input frequency, is safe only when these harmonics can be kept to a very low level, of the order of a few tenths of a percent. In the design of sound-reproducing systems for the home this is never feasible, and limits of 10 or 20 percent are frequently met in practice. Furthermore, on systems of limited frequency range, high-frequency distortion may be audible and annoying, although the actual harmonics may be above the cut-off frequency. This type of distortion has more than nullified the advantages of many a good frequency-response characteristic.

Distortion in Power Output Stages

The procedure for rating harmonic distortion has been greatly oversimplified. For instance, when circuits were simpler, class A triodes were the general rule for power output stages. Such tubes produced relatively little distortion until actually overloaded, and the power output was conveniently rated at the level where 5 percent distortion occurred. Furthermore, such tubes had a low plate impedance, which tended to reduce the effects of a changing load impedance such as a loudspeaker.

In the quest for higher power efficiency, various successive stages of development have included the pentode, the class B amplifier, and the beam tube. Some of these tend to produce appreciable distortion, even at levels well below the maximum power output. Also, some have a high plate impedance, which exaggerates the effects of changing load impedance, thus accentuating the distortion caused by output transformers, shunting capacitance, and the normal changes in loudspeaker impedance. Many of these disadvantages can be overcome by the use of inverse feedback at the expense of gain and simplicity.

Many designers have seriously wondered whether these more elaborate output circuits offer any appreciable advantage, economic or otherwise, over the simpler triode systems. Under commonly encountered operating conditions amplifiers with identical distortion ratings may sound entirely different with degrees of actual audible distortion ranging from practically unnoticeable to practically unbearable.

There are many reasons for this. It is desirable, so far as possible, that the amplifier be operated below its overload point. Under these conditions the actual distortion produced bears little relation to the distortion at some particular degree of overloading. In class B systems, as an extreme, the distortion increases rapidly at low levels. Furthermore, in many systems push-pull and feedback circuits are used, which cause the distortion to vary with frequency. Push-pull amplifiers, for example, are often unbalanced at the extreme ends of the audio-frequency range.

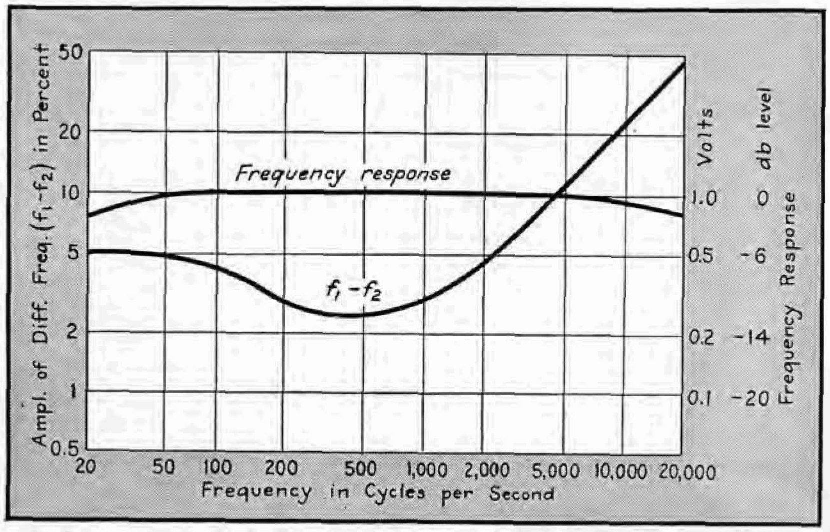

FIG. 2—Frequency-response and intermodulation curves on wide-range amplifier having rough high-frequency reproduction

Some Overloading May be Tolerable

Unfortunately, in equipment designed for home sound reproduction economic considerations limit the power-handling ability of the output stage and the efficiency of the loudspeaker. Consequently the systems are often operated just below the overload point, so that overloading occurs on volume peaks in the music. With good design an amplifier will overload gracefully. The result will be a certain unnatural brightness in the reproduction which may, however, be tolerable for short periods of time, particularly when heard at a high acoustic level, where the ear itself is distorting. The same degree of overloading in a poorly designed amplifier provides a muddy and coarse quality which is infinitely worse to the ear. Harmonic measurements made with single tones give little clue to this difference.

Many writers have pointed out that the intermodulation products and not the harmonics are responsible for the annoying quality when a sound-reproducing system is overloaded. Musical tones contain harmonics at various percentages, sometimes stronger than the fundamental. Adding a small percentage to these harmonics does not in itself produce a serious change in tone quality. When two different tones are passed simultaneously through a distorting amplifier the intermodulation results in sum and difference frequencies which are not harmonically related to the original tones. Some writers have intimated that these components are so low in amplitude as to be negligible, but it is easily demonstrated that this is not the case. We have all heard the soprano solo with flute obbligato marred by the growling of difference frequencies and the symphony orchestra which produces only a confused jumble of sound. Frayne and Scoville[1] showed in a simple mathematical analysis that it is quite possible for such components to be several times the amplitude of the harmonics. They calculated an average ratio of the order of 3.5, which, however, does not hold for conditions more complicated than those which they were considering.

How Intermodulation Occurs

The intermodulation products consist not only of the first-order sum and difference frequencies f_1+f_2 and f_1-f_2 (where f_1 and f_2 are the two fundamentals) but also the second-order terms $2f_1+f_2$, $2f_1-f_2$, f_1+2f_2, f_1-2f_2, and higher-order beats. None of these are harmonically related to the original components in the signal except by accident, hence the harsh discordance characteristic of certain types of so-called harmonic distortion. When the large number of tones involved in the reproduction of a symphony orchestra is considered and it is realized that every tone will intermodulate with every other tone, causing a series of sum and difference frequencies, there is little reason to wonder what causes the blurred effect characteristic of some amplifiers and loud speakers. These effects are what the average listener means by the word distortion.

Previous Intermodulation Research

In certain branches of audio-frequency engineering the presence of intermodulation has produced such serious results as to necessitate more investigation than usual. Frayne and Scoville[1] described an intermodulation test for use in connection with variable-density film recording. Hilliard[2], working in the same field, pointed out the advantages of a similar technique for measuring the performance of amplifiers, radio transmitters, and other systems. He observed that, of two systems having the same total harmonic distortion as measured by conventional means, the one with the greater amount of intermodulation provided reproduction which was definitely more objectionable, and he recommended a means of measuring the intermodulation by applying to the amplifier simultaneously a low and a high audio frequency. In Hilliard's system the higher audio frequency is treated as a modulated carrier and its modulation by the lower frequency measured in much the same way as the modulation of a broadcast station. While the amount of equipment required for such measurements is not negligible, Hilliard reported, "by comparison other methods are inadequate and inconvenient, as well as more laborious." Hilliard considered that the intermodulation had to be less than 2 percent to be unnoticeable to the ear.

In discussing the Hilliard paper, B. F. Meissner[3] pointed out that he had used the two-sine-wave method of test in his development work on electronic musical instruments, analyzing the output with a General Radio wave analyzer, and considered this "the ideal distortion-measuring system, since it measures directly what the ear itself hears as the objectionable element in sound reproduction." Lewis and Hunt[4], in connection with their investigation of tracing distortion in phonograph recording, recognized the importance of the intermodulation components. Their analysis includes re-recording, which is customarily used in the production of vertically cut records in order to minimize tracing distortion, which on this type of record consists mostly of even harmonics and first-order intermodulation products.

At a somewhat earlier date Harries[5] in England used intermodulation measurements to demonstrate the advantages of the so-called Harries valve over the then-current pentodes. Earlier references will be found in European publications, particularly German.[6,7,8,9] Of these Janovsky[8], as early as 1929, performed certain experiments to determine which of the various intermodulation products were most noticeable.

Conventional Distortion Measurements

Analyzers and distortion meters have been developed to a point where harmonic distortion can be measured to 0.1 percent. There are also numerous oscillators available which provide a sufficiently pure signal for these tests. However, intermodulation measurements have not been widely adopted, presumably because of the equipment required, the complexity of the measurements, and the large number of components to be measured and evaluated.

There are many applications where harmonic or distortion-meter measurements alone are inadequate. Home-type sound-reproducing equipment generally operates at distortion levels such that serious intermodulation may be present, and this intermodulation does not have, excepting under a specific set of conditions, a fixed relation to the harmonics. A sharp high-frequency cut-off characteristic will render harmonic measurements useless in the upper octave of the frequency range, yet

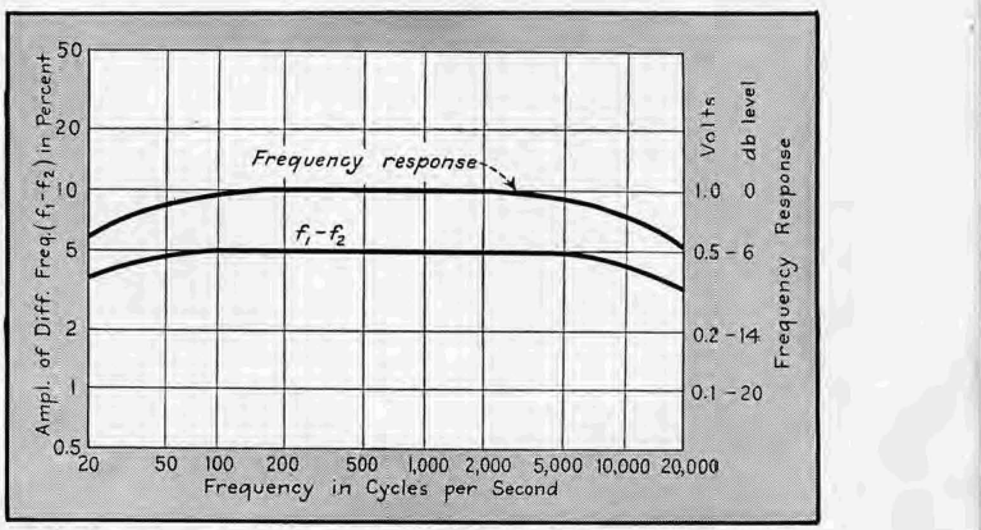

FIG. 3—Frequency-response and intermodulation curves on wide-range amplifier having clean high-frequency reproduction

intermodulation is frequently at its worst in this range. High noise levels, hum, etc., also encountered in low-priced equipment, affect distortion-meter measurements.

Double-Beat Oscillator

A development program had been planned covering complete methods for intermodulation measurements, with the hope of investigating also the relation of such measurements to distortion as judged by the average listener. Like many other programs, this has had to wait for the war, but one instrument has been developed which has proven unusually satisfactory in such applications. This is the fundamental instrument needed for convenient and accurate intermodulation measurements—namely, the source for producing two tones free from harmonics and intermodulation. The new instrument is called the double-beat oscillator and is shown diagrammatically in Fig. 1. Where a standard beat oscillator includes two high-frequency oscillators, this has three. The outputs may be heterodyned in various combinations so as to provide (I) a single variable output frequency, (II) two variable output frequencies having a constant sum or a constant difference, or (III) two independently variable output frequencies. The instrument also includes mixing controls for adjusting the relative amplitudes of the two output frequencies, as well as usual output circuits for varying the total output over wide ranges.

Uses for Double-Beat Oscillators

Such an oscillator may be used in many ways. It will do anything a standard beat oscillator will do, and in addition will provide two simultaneous output frequencies, either one of which may be varied, to allow any sort of intermodulation measurement. Since it can be set so that the two output frequencies, while varied, maintain a fixed sum or difference, the measurement of first-order intermodulation products is greatly simplified and facilitated. In fact, it becomes as easy as running a response curve, since the analyzer tuning may remain fixed, or a simple fixed tuned indicator may be used. This is a tremendous advantage when large numbers of laboratory measurements must be made or for production or routine maintenance checking. The double-beat oscillator is less expensive than two beat-frequency oscillators and much simpler to use because of the constant sum or difference feature and the mixer circuits.

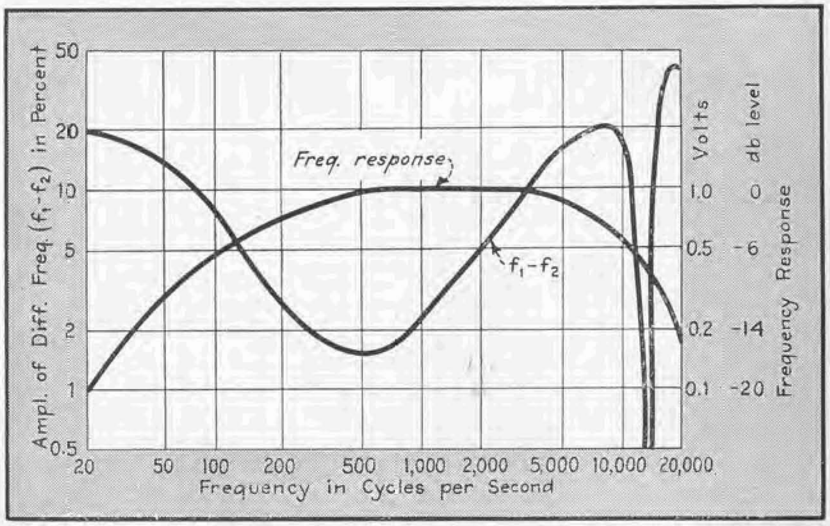

FIG. 4—Frequency-response and intermodulation curves on poor-quality amplifier

Measuring Procedure

Figures 2, 3, 4, and 5 show samples of results obtained with the double-beat oscillator in measuring the difference frequency generated in various types of amplifiers and sound-reproducing systems. The curves were taken as follows: Above 200 cycles a constant difference frequency of 100 cycles was used, the amplitude of the difference frequency being plotted in terms of the higher of the two input frequencies. Since the double-beat oscillator provides two output frequencies with a fixed difference, the analyzer was left tuned and the portion of the curve above 200 cycles obtained by merely varying the main oscillator control. Below 400 cycles the same procedure was used, but with a fixed difference frequency of 500 cycles, the curve being plotted in terms of the lower of the two input frequencies. If the characteristics of the reproducing system are fairly flat between 100 and 500 cycles the curves will overlap almost exactly in this region, thus forming, in effect, a continuous curve.

The main justification for this procedure is purely practical and arbitrary. It is a simple means of obtaining a continuous curve showing first-order intermodulation as a function of the controlling frequency (which is generally the lower of the two frequencies in the low-frequency region and the higher of the two frequencies in the high-frequency region). Such curves on amplifiers producing strong first-order intermodulation check far better with audible estimates than any other simple distortion curves that we have found to date. Janovsky considered the difference tone as the most serious component in this annoying type of distortion.

A similar curve can be obtained without shifting the difference frequency, but two peaks will appear when the fundamentals equal that frequency. This is no disadvantage for routine and production testing, since the difference frequency may be chosen so that it lies in a part of the range where distortion is ordinarily small—for instance, around 400 or 500 cycles.

Analysis of Sample Results

Figure 2 shows an amplifier characterized by a good frequency-response characteristic, but a rough and annoying quality in the high-frequency reproduction. The rise in the difference tone at high frequencies shows the reason. Judged by its frequency-response curve, this is an extremely fine amplifier. On actual listening tests it performed very poorly. The intermodulation characteristic shows at least one very good reason.

Figure 3 shows an amplifier with inferior frequency response to Fig. 2. If only the frequency-response curves were available, one might conclude that the cleaner reproduction of the amplifier shown in Fig. 3 was a result of greater attenuation of high frequencies. Actually, over the important region up to 10,000 cycles the difference between the two amplifiers in this respect would never be noticeable, and the amplifier of Fig. 3 is characterized by unusually clean, crisp, full-range reproduction. The intermodulation curve shows one reason for this. Although ordinary frequency-response and distortion measurements indicate that the amplifier of Fig. 2 is better than that of Fig. 3, listening tests definitely indicate the opposite.

Figure 4 shows another case involving an amplifier with rather poor frequency characteristics, also characterized by harsh reproduction which many designers have attempted to avoid by reducing the high-frequency response. This particular amplifier has bad intermodulation at both the low- and high-frequency ends. Also, in the high-frequency region there is one point where the intermodulation cancels out exactly, which indicates the risk in making intermodulation measurements at only a few frequencies.

Figure 5 shows the actual voltage across the voice coil in a loud speaker - amplifier combination. When operated into its rated load impedance the amplifier is satisfactory, providing less than 2 percent intermodulation. Because of the high output impedance of the amplifier and the variation in the loud speaker impedance with frequency, the intermodulation curve shows sharp rises at the low- and high-frequency regions. While either the amplifier or the loud speaker, when checked alone by conventional methods, seems satisfactory, the combination of the two is definitely not.

Second-Order Products

The foregoing curves show only first-order intermodulation, which may not always be the controlling factor. It is realized that under certain conditions, and particularly in highly balanced circuits, the first-order intermodulation products will tend to cancel and the second-order intermodulation products become the important factors in the audible distortion. It is, of course, possible to build an oscillator having an output such that one or more of the second-order intermodulation products can be kept constant, but this has not seemed warranted so far.

As a practical matter, many push-pull amplifiers do not seem to be as well balanced as might be assumed, particularly at the extremes of the frequency range, so that the first-order intermodulation products are as strong as or stronger than the higher-order products. (The amplifier shown in Fig. 4 is push-pull.) Harries[5] reported that on symmetrical overloading the first-order intermodulation products rose to a maximum as the overloading increased and then fell off as the second-order intermodulation products rose. His observations were on single output tubes having an S-shaped amplitude characteristic, but also seem to apply to many actual push-pull amplifiers. Under these conditions the distortion is generally serious before the first-order intermodulation has reached its maximum.

Conclusions

We have used this double-beat oscillator only on a few special applications, but it has proved so satisfactory and convenient that we feel that there may be a real demand for such instruments in the field. It is mentioned not as a cure-all, but as a further step in the design of audio-frequency measuring equipment in an effort to obtain results which correlate better with listening tests.

The writer will be very glad to receive comments and suggestions from other engineers who have used two-frequency measurements, or who have devised equipment for making such measurements.

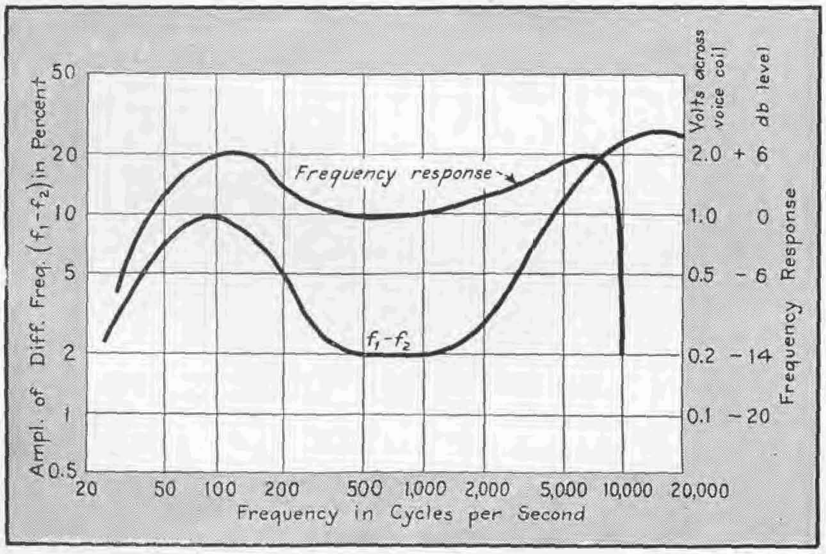

FIG. 5—Frequency-response and intermodulation curves showing effects on amplifier loaded by loudspeaker and whistle filter

REFERENCES

(1) Frayne, J. G., and Scoville, R. R., Variable-Density Recording, *Journal of the Society of Motion Picture Engineers*, **32**, p. 648-673, June, 1939.

(2) Hilliard, John K., Distortion Tests by the Intermodulation Method, *Proc. I.R.E.*, **29**, No. 12, p. 614-620, Dec., 1941.

(3) Meissner, Benjamin F., Discussion on "Distortion Tests by the Intermodulation Method," *Proc. I.R.E.*, **30**, No. 9, p 429, Sept., 1942.

(4) Lewis, W. D., and Hunt, F. V., A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records *Journal of the Acoustical Society of America*, **12**, No. 3, p. 348-365, Jan 1941.

(5) Harries, J. H. Owen, Amplitude Distortion, *The Wireless Engineer*, **14**, No. 161 p. 63-72, Feb., 1937.

(6) A Narath, The Double-Note Method of Distortion Measurement and its Application to Sound Films (including a Survey of the Various Types of Distortion in the Photographic Processes of Sound-on-Film Work), *Telefunken-Zeit.*, **17**, No. 73, p. 57-68, July, 1936.

(7) Wilhelm, K., and Kettel, E., A New Method for the Representation and Measurement of Non-Linear Distortions (using a Cathode-Ray Oscillograph and Two Modulating Frequencies), *Telefunken-Röhre*, No 6, p. 24-35, March, 1936.

(8) Janovsky, W., Über die Hörbarkeit von Verzerrungen (The Audibility of Distortion), *E.N.T.*, **6**, p. 421-439, Nov., 1929.

(9) von Ardenne, Manfred, Rectification as a Criterion of Distortion in Amplifiers *Experimental Wireless and the Wireless Engineer*, **V**, No. 53, p. 52-55, Feb., 1928.